

file  
0590  
3/26 Jumbo

# 3



OIPE

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/989,994

DATE: 11/26/2002 P6  
TIME: 13:52:59

Input Set : A:\seqlist.txt

Output Set: N:\CRF4\11262002\I989994.raw

```
3 <110> APPLICANT: LIU, Qiang
5 <120> TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
6   TRIPLETS BY ZINC FINGERS
8 <130> FILE REFERENCE: 8325-0011.20 / S11-US2
10 <140> CURRENT APPLICATION NUMBER: 09/989,994
11 <141> CURRENT FILING DATE: 2001-11-20
13 <160> NUMBER OF SEQ ID NOS: 4085
15 <170> SOFTWARE: PatentIn Ver. 2.0
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 25
19 <212> TYPE: PRT
20 <213> ORGANISM: Artificial Sequence
22 <220> FEATURE:
23 <223> OTHER INFORMATION: Description of Artificial Sequence: exemplary motif
25 <220> FEATURE:
26 <221> NAME/KEY: MOD_RES
27 <222> LOCATION: (2)..(5)
28 <223> OTHER INFORMATION: this region may encompass two to four residues
29   consisting of any amino acid
31 <220> FEATURE:
32 <221> NAME/KEY: MOD_RES
33 <222> LOCATION: (7)..(18)
34 <223> OTHER INFORMATION: any amino acid
36 <220> FEATURE:
37 <221> NAME/KEY: MOD_RES
38 <222> LOCATION: (20)..(24)
39 <223> OTHER INFORMATION: this region may encompass three to five residues
40   consisting of any amino acid
42 <400> SEQUENCE: 1
W--> 43 Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
44   1           5           10          15
W--> 46 Xaa Xaa His Xaa Xaa Xaa Xaa His
47           20           25
50 <210> SEQ ID NO: 2
51 <211> LENGTH: 5
52 <212> TYPE: PRT
53 <213> ORGANISM: Artificial Sequence
55 <220> FEATURE:
56 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide linker
58 <400> SEQUENCE: 2
59 Thr Gly Glu Lys Pro
60   1           5
63 <210> SEQ ID NO: 3
```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/989,994

DATE: 11/26/2002

TIME: 13:52:59

Input Set : A:\seqlist.txt

Output Set: N:\CRF4\11262002\I989994.raw

```

64 <211> LENGTH: 5
65 <212> TYPE: PRT
66 <213> ORGANISM: Artificial Sequence
68 <220> FEATURE:
69 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide linker
71 <400> SEQUENCE: 3
72 Gly Gly Gly Gly Ser
73   1           5
76 <210> SEQ ID NO: 4
77 <211> LENGTH: 8
78 <212> TYPE: PRT
79 <213> ORGANISM: Artificial Sequence
81 <220> FEATURE:
82 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide linker
84 <400> SEQUENCE: 4
85 Gly Gly Arg Arg Gly Gly Gly Ser
86   1           5
89 <210> SEQ ID NO: 5
90 <211> LENGTH: 9
91 <212> TYPE: PRT
92 <213> ORGANISM: Artificial Sequence
94 <220> FEATURE:
95 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide linker
97 <400> SEQUENCE: 5
98 Leu Arg Gln Arg Asp Gly Glu Arg Pro
99   1           5
102 <210> SEQ ID NO: 6
103 <211> LENGTH: 12
104 <212> TYPE: PRT
105 <213> ORGANISM: Artificial Sequence
107 <220> FEATURE:
108 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide linker
110 <400> SEQUENCE: 6
111 Leu Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro
112   1           5           10
115 <210> SEQ ID NO: 7
116 <211> LENGTH: 16
117 <212> TYPE: PRT
118 <213> ORGANISM: Artificial Sequence
120 <220> FEATURE:
121 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide linker
123 <400> SEQUENCE: 7
124 Leu Arg Gln Lys Asp Gly Gly Gly Ser Gly Gly Gly Ser Glu Arg Pro
125   1           5           10           15
128 <210> SEQ ID NO: 8
129 <211> LENGTH: 25
130 <212> TYPE: PRT
131 <213> ORGANISM: Artificial Sequence
133 <220> FEATURE:

```

## RAW SEQUENCE LISTING

DATE: 11/26/2002

PATENT APPLICATION: US/09/989,994

TIME: 13:52:59

Input Set : A:\seqlist.txt

Output Set: N:\CRF4\11262002\I989994.raw

```

134 <223> OTHER INFORMATION: Description of Artificial Sequence: exemplary motif
136 <220> FEATURE:
137 <221> NAME/KEY: MOD_RES
138 <222> LOCATION: (2)..(5)
139 <223> OTHER INFORMATION: this region may encompass two to four residues
140     consisting of any amino acid
142 <220> FEATURE:
143 <221> NAME/KEY: MOD_RES
144 <222> LOCATION: (7)..(18)
145 <223> OTHER INFORMATION: any amino acid
147 <220> FEATURE:
148 <221> NAME/KEY: MOD_RES
149 <222> LOCATION: (20)..(24)
150 <223> OTHER INFORMATION: this region may encompass three to five residues
151     consisting of any amino acid
153 <400> SEQUENCE: 8
W--> 154 Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
155      1             5             10             15
W--> 157 Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His
158           20           25
161 <210> SEQ ID NO: 9
162 <211> LENGTH: 30
163 <212> TYPE: PRT
164 <213> ORGANISM: Artificial Sequence
166 <220> FEATURE:
167 <223> OTHER INFORMATION: Description of Artificial Sequence: mouse transcription
168     factor Zif268
170 <400> SEQUENCE: 9
171 Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp
172   1             5             10             15
174 Glu Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro
175           20           25           30
178 <210> SEQ ID NO: 10
179 <211> LENGTH: 28
180 <212> TYPE: PRT
181 <213> ORGANISM: Artificial Sequence
183 <220> FEATURE:
184 <223> OTHER INFORMATION: Description of Artificial Sequence: mouse transcription
185     factor Zif268
187 <400> SEQUENCE: 10
188 Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu
189   1             5             10             15
191 Thr Thr His Ile Arg Thr His Thr Gly Glu Lys Pro
192           20           25
195 <210> SEQ ID NO: 11
196 <211> LENGTH: 27
197 <212> TYPE: PRT
198 <213> ORGANISM: Artificial Sequence
200 <220> FEATURE:

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/989,994

DATE: 11/26/2002

TIME: 13:52:59

Input Set : A:\seqlist.txt

Output Set: N:\CRF4\11262002\I989994.raw

201 <223> OTHER INFORMATION: Description of Artificial Sequence: mouse transcription  
 202 factor Zif268  
 204 <400> SEQUENCE: 11  
 205 Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg  
 206 1 5 10 15  
 208 Lys Arg His Thr Lys Ile His Leu Arg Gln Lys  
 209 20 25  
 212 <210> SEQ ID NO: 12  
 213 <211> LENGTH: 9  
 214 <212> TYPE: DNA  
 215 <213> ORGANISM: Artificial Sequence  
 217 <220> FEATURE:  
 218 <223> OTHER INFORMATION: Description of Artificial Sequence: target DNA  
 220 <400> SEQUENCE: 12  
 221 gcgtgggcg 9  
 224 <210> SEQ ID NO: 13  
 225 <211> LENGTH: 94  
 226 <212> TYPE: PRT  
 227 <213> ORGANISM: Artificial Sequence  
 229 <220> FEATURE:  
 230 <223> OTHER INFORMATION: Description of Artificial Sequence: Sp-1  
 231 transcription factor  
 233 <400> SEQUENCE: 13  
 234 Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly Lys  
 235 1 5 10 15  
 237 Val Tyr Gly Lys Thr Ser His Leu Arg Ala His Leu Arg Trp His Thr  
 238 20 25 30  
 240 Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg Phe  
 241 35 40 45  
 243 Thr Arg Ser Asp Glu Leu Gln Arg His Lys Arg Thr His Thr Gly Glu  
 244 50 55 60  
 246 Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Arg Ser Asp  
 247 65 70 75 80  
 249 His Leu Ser Lys His Ile Lys Thr His Gln Asn Lys Lys Gly  
 250 85 90  
 253 <210> SEQ ID NO: 14  
 254 <211> LENGTH: 9  
 255 <212> TYPE: DNA  
 256 <213> ORGANISM: Artificial Sequence  
 258 <220> FEATURE:  
 259 <223> OTHER INFORMATION: Description of Artificial Sequence: target DNA  
 261 <400> SEQUENCE: 14  
 262 ggggcgggg 9  
 265 <210> SEQ ID NO: 15  
 266 <211> LENGTH: 100  
 267 <212> TYPE: PRT  
 268 <213> ORGANISM: Artificial Sequence  
 270 <220> FEATURE:  
 271 <223> OTHER INFORMATION: Description of Artificial Sequence: Sp-1 consensus

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/989,994

DATE: 11/26/2002

TIME: 13:52:59

Input Set : A:\seqlist.txt

Output Set: N:\CRF4\11262002\I989994.raw

```

272      sequence
274 <400> SEQUENCE: 15.
275 Met Glu Lys Leu Arg Asn Gly Ser Gly Asp Pro Gly Lys Lys Lys Gln
276   1              5              10              15
278 His Ala Cys Pro Glu Cys Gly Lys Ser Phe Ser Lys Ser Ser His Leu
279              20              25              30
281 Arg Ala His Gln Arg Thr His Thr Gly Glu Arg Pro Tyr Lys Cys Pro
282              35              40              45
284 Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Gln Arg His Gln
285              50              55              60
287 Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
288   65              70              75              80
290 Ser Phe Ser Arg Ser Asp His Leu Ser Lys His Gln Arg Thr His Gln
291              85              90              95
293 Asn Lys Lys Gly
294              100
297 <210> SEQ ID NO: 16
298 <211> LENGTH: 9
299 <212> TYPE: DNA
300 <213> ORGANISM: Artificial Sequence
302 <220> FEATURE:
303 <223> OTHER INFORMATION: Description of Artificial Sequence: Sp-1 consensus
304      sequence
306 <400> SEQUENCE: 16
307 ggggcggggg 9
310 <210> SEQ ID NO: 17
311 <211> LENGTH: 9
312 <212> TYPE: DNA
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <223> OTHER INFORMATION: Description of Artificial Sequence: example target
317      DNA
319 <400> SEQUENCE: 17
320 gcggggggcg 9
323 <210> SEQ ID NO: 18
324 <211> LENGTH: 9
325 <212> TYPE: DNA
326 <213> ORGANISM: Artificial Sequence
328 <220> FEATURE:
329 <223> OTHER INFORMATION: Description of Artificial Sequence: example target
330      DNA
332 <400> SEQUENCE: 18
333 gcggggggcg 9
336 <210> SEQ ID NO: 19
337 <211> LENGTH: 9
338 <212> TYPE: DNA
339 <213> ORGANISM: Artificial Sequence
341 <220> FEATURE:
342 <223> OTHER INFORMATION: Description of Artificial Sequence: example target

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/989,994

DATE: 11/26/2002  
TIME: 13:53:00

Input Set : A:\seqlist.txt  
Output Set: N:\CRF4\11262002\I989994.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 2,3,4,5,7,8,9,10,11,12,13,14,15,16,17,18,20,21,22,23,24  
Seq#:8; Xaa Pos. 2,3,4,5,7,8,9,10,11,12,13,14,15,16,17,18,20,21,22,23,24  
Seq#:4013; Xaa Pos. 2,5,6  
Seq#:4014; Xaa Pos. 2,5,6  
Seq#:4015; Xaa Pos. 2,5,6  
Seq#:4016; Xaa Pos. 2,5,6  
Seq#:4017; Xaa Pos. 2,5,6  
Seq#:4018; Xaa Pos. 2,5,6  
Seq#:4019; Xaa Pos. 2,5,6  
Seq#:4020; Xaa Pos. 2,5,6  
Seq#:4021; Xaa Pos. 2,5,6  
Seq#:4022; Xaa Pos. 2,5,6  
Seq#:4023; Xaa Pos. 2,5,6  
Seq#:4024; Xaa Pos. 2,5,6  
Seq#:4025; Xaa Pos. 2,5,6  
Seq#:4026; Xaa Pos. 2,5,6  
Seq#:4027; Xaa Pos. 2,5,6  
Seq#:4028; Xaa Pos. 2,5,6  
Seq#:4029; Xaa Pos. 2,5,6  
Seq#:4030; Xaa Pos. 2,5,6  
Seq#:4031; Xaa Pos. 2,5,6  
Seq#:4032; Xaa Pos. 2,5,6  
Seq#:4033; Xaa Pos. 2,5,6  
Seq#:4034; Xaa Pos. 2,5,6  
Seq#:4035; Xaa Pos. 2,5,6  
Seq#:4036; Xaa Pos. 2,5,6  
Seq#:4037; Xaa Pos. 2,5,6  
Seq#:4038; Xaa Pos. 2,5,6  
Seq#:4039; Xaa Pos. 2,5,6  
Seq#:4040; Xaa Pos. 2,5,6  
Seq#:4041; Xaa Pos. 2,5,6  
Seq#:4042; Xaa Pos. 2,5,6  
Seq#:4043; Xaa Pos. 2,5,6  
Seq#:4044; Xaa Pos. 2,5,6  
Seq#:4045; Xaa Pos. 2,5,6  
Seq#:4046; Xaa Pos. 2,5,6  
Seq#:4047; Xaa Pos. 2,5,6  
Seq#:4048; Xaa Pos. 2,5,6  
Seq#:4049; Xaa Pos. 2,5,6  
Seq#:4050; Xaa Pos. 2,5,6  
Seq#:4051; Xaa Pos. 2,5,6  
Seq#:4052; Xaa Pos. 2,5,6  
Seq#:4053; Xaa Pos. 2,5,6  
Seq#:4054; Xaa Pos. 2,5,6

**RAW SEQUENCE LISTING ERROR SUMMARY**  
PATENT APPLICATION: **US/09/989,994**DATE: 11/26/2002  
TIME: 13:53:00Input Set : **A:\seqlist.txt**  
Output Set: **N:\CRF4\11262002\I989994.raw**

Seq#:4055; Xaa Pos. 2,5,6  
Seq#:4056; Xaa Pos. 2,5,6  
Seq#:4057; Xaa Pos. 2,5,6  
Seq#:4058; Xaa Pos. 2,5,6  
Seq#:4059; Xaa Pos. 2,5,6  
Seq#:4060; Xaa Pos. 2,5,6  
Seq#:4061; Xaa Pos. 2,5,6  
Seq#:4062; Xaa Pos. 2,5,6  
Seq#:4063; Xaa Pos. 2,5,6  
Seq#:4064; Xaa Pos. 2,5,6  
Seq#:4065; Xaa Pos. 2,5,6  
Seq#:4066; Xaa Pos. 2,5,6  
Seq#:4067; Xaa Pos. 2,5,6  
Seq#:4068; Xaa Pos. 2,5,6  
Seq#:4069; Xaa Pos. 2,5,6  
Seq#:4070; Xaa Pos. 2,5,6  
Seq#:4071; Xaa Pos. 2,5,6  
Seq#:4072; Xaa Pos. 2,5,6  
Seq#:4073; Xaa Pos. 2,5,6  
Seq#:4074; Xaa Pos. 2,5,6  
Seq#:4075; Xaa Pos. 2,5,6  
Seq#:4076; Xaa Pos. 2,5,6  
Seq#:4077; Xaa Pos. 2,5,6  
Seq#:4078; Xaa Pos. 2,5,6  
Seq#:4079; Xaa Pos. 2,5,6  
Seq#:4080; Xaa Pos. 2,5,6  
Seq#:4081; Xaa Pos. 2,5,6  
Seq#:4082; Xaa Pos. 2,5,6  
Seq#:4083; Xaa Pos. 2,5,6  
Seq#:4085; Xaa Pos. 4,6

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/989,994

DATE: 11/26/2002

TIME: 13:53:00

Input Set : A:\seqlist.txt

Output Set: N:\CRF4\11262002\I989994.raw

L:53292 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4057 after pos.:0

L:53316 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4058 after pos.:0



US 0998999409P1



Creation date: 17-07-2003  
Indexing Officer: TDANG2 - TRAI DANG  
Team: OIPEBackFileIndexing  
Dossier: 09989994

Legal Date: 10-03-2003

No.	Doccode	Number of pages
1	IDS	12

Total number of pages: 12

Remarks:

Order of re-scan issued on .....